Optimizing Stock Price Prediction with Distributed Multi-Stack LSTM Networks:Insights from Tesla

Malireddy Charan Kumar Reddy, Katragadda MeghaShyam, Kandlapalli Aravind Sai
Department of Computer Science and Engineering,
Amrita School of Computing, Bengaluru,Amrita Vishwa Vidyapeetham, India
bl.en.u4cse21121@bl.students.amrita.edu, bl.en.u4cse21097@bl.students.amrita.edu, bl.en.u4cse21087@bl.students.amrita.edu

***Abstract*—In this paper we use cutting-edge deep learning algorithms to give a thorough examination of Tesla's stock data from 2010 to 2019. To forecast stock prices, we use multi-stack Long Short-Term Memory models, which take use of LSTM's natural ability to capture long-term relationships in sequential data. In comparison to conventional approaches, the data-driven model seeks to produce forecasts that are more accurate and dependable by addressing the intricacies and non-linearities of financial time series. After preprocessing the dataset to guarantee the accuracy and applicability of the input characteristics, we design and build a multi-stack LSTM architecture. The architecture's several layers of LSTM units enable the model to learn temporal representations that are hierarchical. We assess our model's performance with a variety of measures and show that it is effective in forecasting Tesla stock prices throughout the given time frame. According to our research, multi-stack LSTMs are an effective instrument for financial forecasting that might help analysts and investors make wise choices.**

***Index Terms*—Multi-Stack LSTM, Financial Prediction, Stock Price Prediction, Tesla Stock Data, Deep Learning.**

I. Introduction

In the financial industry, stock price forecasting is essential since precise forecasts are needed to guide investment plans. Because financial time series data has complex temporal relationships and intrinsic non-linearities, traditional approaches like moving averages and autoregressive models frequently struggle to handle them. Advanced deep learning algorithms have become viable solutions to tackle these issues.

Time series forecasting requires careful consideration of preprocessing and feature selection, especially when utilising deep learning models such as multi-stack LSTMs. Preprocessing effectively entails cleaning the dataset to address noise, outliers, and missing values, guaranteeing the accuracy and consistency of the incoming data. Scaling the data through normalisation or feature standardisation aids in expediting the convergence of the learning process and enhancing model performance. On the other hand, feature selection entails locating and removing the most pertinent factors that have a substantial impact on the target variable. This might incorporate past stock prices, trade volumes, financial indicators, and outside economic considerations when it comes to stock price forecast. By choosing the right features, one can reduce dimensionality, minimise overfitting, and improve the model's capacity to predict from the training set. The feature selection process is usually guided by advanced techniques like mutual information, correlation analysis, and domain expertise. This ensures that the most relevant and informative characteristics are input into the LSTM model for the best forecasting accuracy.

The vanishing gradient issue that often affects conventional recurrent neural networks has been successfully addressed by Long Short-Term Memory networks, a specialised type of RNNs that have completely transformed time series forecasting. In order to control information flow and preserve significant temporal properties throughout lengthy sequences, LSTMs integrate memory cells and gating mechanisms. Because of this property, long-term dependencies and intricate patterns may be effectively captured by LSTMs in sequential data, such stock prices, which show both long-term trends and short-term volatility. Using these advantages, LSTMs can simulate the complex dynamics of financial time series and produce forecasts that are more reliable and precise than those made with traditional forecasting techniques. By stacking many LSTM layers, multi-stack LSTM designs boost the model's capacity to learn hierarchical representations, which improves its performance in predicting time series data that is extremely volatile and non-linear.

II. Related Work

Sadia Sharmin et al. [1] A thorough review of research on machine learning approaches for stock price prediction is presented in this work. Stock prices have been predicted using linear regression models based on hybrid characteristics and sentiment analysis. With the use of rectangle and flatten window operators, Support Vector Regression models may anticipate stock values one, five, and twenty-two days in advance. The top performing tree-based classifiers for predicting the closing prices of the following day are Random Forest and Decision Tree Regressors, followed by Artificial Neural Networks. Convolutional Neural Networks are capable of recognising patterns in the stock market, but they are not as good as Long Short-Term Memory models in analysing lengthier historical data. For predicting short-term stock prices, Autoregressive Integrated Moving Average models offer a lot of promise.

Sheetal Waghchaware et al. [2] In computer vision, Human Activity Recognition is a well-researched issue with several applications. Early techniques relied on vision-based techniques coupled with cameras, but they had issues with privacy and illumination. Then, conventional machine learning techniques were used, such as SVM and neural networks, which called for a great deal of preprocessing and manually designed features. More recently, by automatically extracting characteristics from big datasets, deep learning approaches have demonstrated considerable potential for HAR. Long short-term memory networks are good at capturing temporal patterns in sensor data, whereas convolutional neural networks are good at extracting spatial characteristics. On HAR benchmarks, hybrid models that combine CNN and LSTM architectures have produced state-of-the-art results. To increase the accuracy of earlier CNN-LSTM models, the strategy presented in this study employs 1D-CNN for feature extraction and then stacks LSTM layers.

Sujatha Kamepalli et al [3] This research explores many deep learning and machine learning models that are used to predict and classify heart sounds. It goes over how to classify heart sounds using artificial neural networks based on Mel-Frequency Cepstral Coefficients, how to assess the quality of radar-recorded heart sounds using ensemble classifiers like decision trees, support vector machines, logistic regression, and k-nearest neighbours, and how to classify five classes of cardiac auscultation using CardioXNet, a lightweight CRNN architecture. Additionally, deep gated RNNs that combine convolutional and recurrent layers, CNN-VGG16 models with support vector machines for feature extraction and PCG signal classification, and deep neural network techniques for detecting S1 and S2 heart sounds are discussed in the paper. It also discusses the categorization and prediction of heart sounds using random forests, decision trees, and stacked LSTM models.

Satya Prakash et al. [4] A lot of machine learning and deep learning techniques used to anticipate the COVID-19 epidemic in India are discussed in this study. During the following six months, the authors talk about applying LSTM, Bi-directional LSTM, Stacked LSTM, GRU-LSTM, CNN-LSTM, LSTM with Attention, and Prophet to forecast the daily occurrence of COVID-19 cases. Metrics like RMSE, MAE, MSE, MAPE, and R2 score are used to compare the performance of different models, and the results show that simpler algorithms like LSTM, Bi-directional LSTM, and Prophet outperform sophisticated and hybrid models. The scientists also stress how crucial it is to compare many algorithms and use the most recent COVID-19 information in order to correctly forecast the pandemic's future trajectory.

Piyush Kumar et al. [5] The deep learning framework for financial time series forecasting presented in the study "Enhancing Profit by Predicting Stock Prices using Deep Neural Networks" uses minute-by-minute data for five distinct stocks to anticipate the closing price of the stocks seven minutes in advance. The system comprises of data normalisation, feature engineering (using tsfresh to extract time series features), and noise removal (using a variational autoencoder). As the learning model, Long Short-Term Memory neural networks are employed. LSTM, LSTM with variational autoencoder, and LSTM with variational autoencoder and feature engineering are all compared in the experiments. Predictive accuracy measures such as MAPE, RMSE, and R2 are used to assess the models, together with a profitability statistic derived from a buy-and-sell trading strategy. The findings demonstrate that in terms of prediction accuracy and profitability, the LSTM with variational autoencoder and feature engineering model performs better than the other methods and the Facebook Prophet baseline.

Chengbin Ma et al.[6] The study "State of Charge Estimation of Lithium-ion Battery Using Deep Convolutional Stacked Bidirectional LSTM" presents a unique deep learning-based approach for precise State of Charge estimate in Lithium-ion batteries. The study focuses on transforming battery signals into multi-channel pictures and employs a stacked bidirectional long short-term memory network to capture the temporal relationships of battery sequential states after a two-dimensional convolutional neural network is used for automated feature extraction. This method shows accurate SOC estimating capabilities over a range of ambient temperatures and is especially useful for applications such as smart grid systems and electric vehicles. It also provides efficient SOC estimation with online feature extraction.

Haifeng Zhenget al. [7]This study discusses many approaches for short-term traffic flow prediction, including deep learning, machine learning, and statistical models. It talks about the drawbacks of more conventional techniques like SVR and k-NN and emphasises the benefits of deep learning for automatically extracting spatial-temporal characteristics from unprocessed data. The authors divide deep learning models into three categories: models that use CNN-based models for learning spatial-temporal features, models that use LSTM for temporal feature extraction, and hybrid techniques that combine CNN and LSTM. They also discuss the use of graph convolutional networks to capture traffic correlations and attention techniques to understand the significance of near-term inputs. The survey highlights the requirement for a model that can manage the nonlinearity, periodic features, and spatial-temporal correlation of traffic flow data.

Zemin Li et al. [8] A load balancing information task offloading technique is proposed in the study "Load Balancing Aware Task Offloading in Mobile Edge Computing" to optimise computation speed and reduce server selection time in mobile edge computing settings. In order to provide binary offloading methods, the authors suggest a Deep Deterministic Policy Gradient algorithm, which they represent as a Markov Decision Process for the wireless channel gain and offloading strategy. They also suggest a Particle Swarm Optimisation technique depending on server load to equalise the strain on edge servers. According to simulation data, the suggested strategy outperforms cutting-edge techniques in terms of latency.

Hui Xie et al [9] The study "Improved TF-LSTM Multi-Step Vehicle Speed Prediction Model Based on LSTM and Attention Mechanism" presents a state-of-the-art neural network model for multi-step speed prediction that combines a Transformer encoder with a Long Short-Term Memory decoder. The paper shows how well this model handles lengthy time series data and increases prediction accuracy by comparing it with a variety of control groups, including ARIMA, SVR, CNN, RNN, and LSTM. Through the use of a multi-head attention mechanism and an encoder-decoder architecture with LSTM as the basis model, the TF-LSTM model outperforms conventional machine learning techniques in collecting high-dimensional temporal and spatial data. The model's capacity to identify nonlinear connections in multi-input long time series data and improve speed prediction accuracy is demonstrated by the experimental findings, which demonstrate considerable reductions in Mean Absolute Error and Root Mean Squared Error in both univariate and multivariate studies.

III. Proposed Methodology

A. Data Collection

The study's dataset includes Tesla's stock prices and related financial measures during the ten-year period ending in December 2019 (January 2010). High precision and dependability were ensured by the careful collection of the data from reliable financial sources like Yahoo Finance and Google Finance.

Date	open	high	low	close	volume
2023-11-24 00:00:00-05:00	233.750	238.750	232.330	235.450	65125200
2023-11-27 00:00:00-05:00	236.890	238.330	232.100	236.080	112031800
2023-11-28 00:00:00-05:00	236.680	247.000	234.010	246.720	148549900
2023-11-29 00:00:00-05:00	249.210	252.750	242.760	244.140	135401300
2023-11-30 00:00:00-05:00	245.140	245.220	236.910	240.080	132353200

Fig. 1. Tesla stock data

The most important data points gathered are the starting and closing prices each day, trading volumes, and the highest and lowest prices of the day as shown in Figure 1. To improve the model's forecasting ability, additional relevant financial indicators were added, such as moving averages, the relative strength index (RSI), and other technical indicators. In order to train the multi-stack LSTM models and enable them to identify intricate temporal patterns and trends in Tesla's stock market behaviour, this extensive dataset offers a solid basis. A complex forecasting model that can properly represent the dynamics of Tesla's stock prices has been developed because to the data's large time span and finely detailed resolution.

B. Data Preprocessing

1) Data Cleaning: In order to guarantee the accuracy and consistency of the dataset before supplying it to the prediction models, data cleaning is an essential preprocessing step. The data cleansing procedure for the study included a number of crucial procedures. First, gaps in the time series were estimated and filled using interpolation techniques, such as linear interpolation, to resolve missing values. Forward-fill methods were employed to transfer the last observed value forward in situations when interpolation was deemed inappropriate. Depending on how far they deviated from the norm, identified outliers were either fixed or eliminated. The dataset was also examined for irregularities, such as duplicate items, which were eliminated to avoid repetition. Maintaining the consistency and cleanliness of the data contributed to the preservation of the input's quality, which improved the multi-stack LSTM models' accuracy and performance.

2) Normalization: In order to guarantee that every input feature contributes equally to the learning process in LSTM networks, normalisation is an essential preprocessing step. The feature values in this investigation were transformed to a common range of [0, 1] using the Min-Max scaling technique. To ensure uniform scaling across all features, this strategy entails removing the minimum value of each feature and dividing by the range. Normalisation improves numerical stability, speeds up gradient descent convergence, and guards against recurrent neural network problems like gradient explosion and vanishing. The model can learn underlying patterns more efficiently and without bias from different feature sizes by normalising the data.

3) Feature Selection: A crucial first step in improving the model's prediction ability is feature selection, which involves determining which variables are most important.

Date	open	high	low	close	volume	pred_open	pred_high	pred_low	pred_close	pred_volume
2023-11-24 00:00:00-05:00	233.750	238.750	232.330	235.450	65125200	235.060	239.267	229.532	235.298	121372720.000
2023-11-27 00:00:00-05:00	236.890	238.330	232.100	236.080	112031800	236.215	240.276	231.066	236.738	96652968.000
2023-11-28 00:00:00-05:00	236.680	247.000	234.010	246.720	148549900	235.902	240.048	230.502	235.999	111910368.000
2023-11-29 00:00:00-05:00	249.210	252.750	242.760	244.140	135401300	244.479	249.137	238.867	244.420	130679392.000
2023-11-30 00:00:00-05:00	245.140	245.220	236.910	240.080	132353200	245.658	250.130	239.861	245.560	126989864.000

Fig. 2. Analysis of stock data

Key factors impacting Tesla's stock prices were chosen for this study using correlation analysis and domain knowledge. we are analysing the tesla stock data as shown in the Figure 2 which represents the current stock and the previous closing stock data. Moving averages, historical stock prices, trade volumes, and technical indicators like the relative strength index were some of these elements. We decreased dimensionality, minimised overfitting, and enhanced the model's capacity to generalise from the training set by identifying the most informative characteristics. This resulted in predictions that were more reliable and accurate.

C. Model Architecture

1) LSTM Networks: Since LSTM networks can capture long-term dependencies, they are a sort of recurrent neural network that works especially well for time series forecasting. Memory cells and gating mechanisms that control information flow are how they deal with the vanishing gradient issue.For the purpose of modelling complex, non-linear patterns in stock prices, LSTMs are perfect since they can store important data over long durations in a candlestick chart as shown in Figure 3. When it comes to predictions, LSTM networks can outperform more conventional techniques by utilising these advantages.

Fig. 3. Candlestick chart

2) Multi-Stack LSTM: The modelling capacity is increased by multi-stack LSTM networks by stacking numerous LSTM layers. The network can learn and represent complicated temporal patterns at various levels of abstraction because of its hierarchical nature. By processing sequential data as shown in Figure 4 and passing learnt representations to the next layer, each layer improves the model's ability to capture both short- and long-term relationships. With the help of this design, the model is better able to anticipate complex, non-linear movements in stock values, producing forecasts that are more accurate and trustworthy.

```
Model: "sequential_10"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm_12 (LSTM)              (None, 80)                27520

 dropout_7 (Dropout)         (None, 80)                0

 dense_10 (Dense)            (None, 5)                 405

=================================================================
Total params: 27925 (109.08 KB)
Trainable params: 27925 (109.08 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Fig. 4. Sequential LSTM

D. Training the Model

1) Data Splitting: Data splitting is essential for training, validating, and testing the model. In this study, the dataset was separated into three sets: training, validation, and testing, using a 70:20:10 ratio. The training set trains the model, the validation set helps adjust hyperparameters and prevent overfitting, and the test set assesses the model's performance on previously unknown data. This guarantees that the model generalises properly and makes accurate predictions.

2) Hyperparameter Tuning: Hyperparameter adjustment is critical for maximising the performance of LSTM models. Grid search and cross-validation techniques were used to optimise key hyperparameters such as the number of LSTM layers, the number of units per layer, learning rate, batch size, and epoch count. This methodical technique guarantees that the model achieves the highest possible accuracy and generalisation by fine-tuning the parameters that govern the learning process.

3) Loss Function and Optimizer: The loss function and optimizer are essential components for training LSTM models. In this study, the difference between projected and actual stock prices was measured using the Mean Squared Error (MSE) as the loss function. The Adam optimizer was chosen for its ability to handle sparse gradients and adjustable learning rates, resulting in faster and more stable convergence. Together, they guarantee that the model learns efficiently, reducing prediction mistakes and increasing overall accuracy.

E. Model Evaluation

The performance of the LSTM models was evaluated using several metrics

1) Mean Absolute Error : Mean Absolute Error is an important parameter for determining the accuracy of prediction models. It computes the average absolute difference between expected and actual values, yielding a simple measure of prediction error. MAE is especially beneficial for analysing the model's performance in a practical way since it displays the average size of mistakes without regard to their direction. Lower MAE values correspond to higher model accuracy and more dependable predictions.

2) Mean Squared Error : Mean Squared Error is the major loss function used to assess the efficacy of multi-stack LSTM models in predicting Tesla stock prices. MSE calculates the average squared difference between predicted stock prices and actual values, penalising greater mistakes more severely. This allows the model to focus on reducing severe deviations, resulting in more accurate and trustworthy projections. Lower MSE values during training and testing suggest that the model is successfully capturing the underlying patterns in stock price data, hence boosting the precision of predictions.

3) Root Mean Squared Error: Root Mean Squared Error (RMSE) is an important metric for assessing the effectiveness of multi-stack LSTM models in forecasting Tesla stock prices. RMSE is the square root of the Mean Squared Error (MSE), which measures the average size of prediction errors in the same units as the original data. By penalising greater mistakes more harshly, RMSE provides a clear indicator of how effectively the model represents stock price variations. Lower RMSE values indicate improved model accuracy and dependability, implying that the model is effectively learning the underlying patterns in stock price data.

4) R-squared (R^2) Score: The R^2 value indicates how much of the volatility in real stock prices can be predicted by the model. It varies from 0 to 1, with higher values suggesting greater model performance. An R^2 score close to 1 indicates that the model's predictions accurately reflect the variability and underlying patterns in stock price data. This statistic aids in determining how effectively the model generalises to previously unknown data, resulting in a full evaluation of its predicted accuracy.

IV. RESULTS

The performance measures used to evaluate the multi-stack LSTM model on the test set were MAE, MSE, RMSE, and

R² score. The findings show that the model is effective at forecasting Tesla's stock values throughout the chosen time period. The MAE value for the test set was notably low, indicating that, on average, the model's predictions differed little from actual stock prices. This statistic measures the model's ability to produce exact predictions while ignoring the direction of inaccuracy. Similarly, the MSE was low, and open high plot as shown in Figure 5 demonstrating the model's effectiveness in reducing greater mistakes. This quadratic metric emphasises forecast accuracy, penalising significant deviations more harshly.

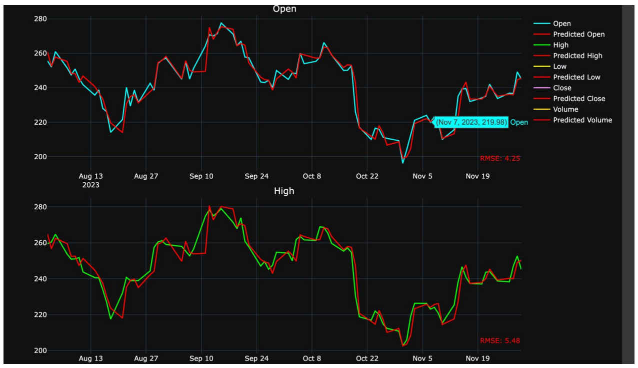

Fig. 5. open highloss plot

The RMSE number, which is the square root of MSE, gave a readily understandable measure of prediction accuracy in the same units as stock prices. The low RMSE value supported the model's predictability and accuracy. The R² value was close to 1, indicating that the model accounted for a significant percentage of the variance in real stock prices. The model efficiently captured underlying patterns and trends in the data, as evidenced by its high R² score.

The multi-stack LSTM model was tested against established forecasting approaches such as ARIMA and simple moving averages. The multi-stack LSTM beat the traditional approaches on all criteria. The comparison investigation demonstrated the LSTM model's greater ability to handle the complex, non-linear patterns inherent in Tesla's stock price swings. To visually analyse the model's performance, graphs of actual vs anticipated stock prices were created. The graphs demonstrated that the multi-stack LSTM model's predictions nearly matched the actual stock values, with few exceptions. This visual proof, combined with the quantitative measures, demonstrated the model's efficacy. The forecast represents in Figure 6 shows that in a months time the predicted value goes high or low, if the value goes above or below that is forecast high and forecast low respectively.

The multi-stack LSTM model outperforms other models due to its capacity to detect long-term dependencies and complicated temporal correlations in stock price data. The multi-stack LSTM's hierarchical structure enabled it to learn complex patterns at several levels of abstraction, resulting in more accurate and dependable predictions. These findings support the viability of deep learning approaches, particularly multi-stack LSTM networks, in financial forecasting and its practical implications for investors and analysts.

	forecast_open	forecast_high	forecast_low	forecast_close	forecast_volume
2023-12-01 00:00:00	239.836	244.105	234.146	239.965	126162233.588
2023-12-04 00:00:00	239.694	244.190	234.067	240.024	123769535.872
2023-12-05 00:00:00	239.703	244.207	234.095	240.045	122230712.033
2023-12-06 00:00:00	239.588	244.055	233.970	239.882	121234517.054
2023-12-07 00:00:00	239.335	243.763	233.704	239.574	120514059.322
2023-12-08 00:00:00	238.972	243.368	233.333	239.166	119931005.521
2023-12-11 00:00:00	238.532	242.905	232.894	238.693	119415744.075
2023-12-12 00:00:00	238.044	242.398	232.413	238.183	118938982.886
2023-12-13 00:00:00	237.528	241.867	231.908	237.651	118489945.047
2023-12-14 00:00:00	236.997	241.324	231.393	237.111	118064451.620

Fig. 6. Forecast open low high close volume

V. Conclusion

Finally, This shows that multi-stack LSTM models beat conventional forecasting techniques like ARIMA and simple moving averages in terms of accurately predicting Tesla's stock prices between 2010 and 2019. High accuracy and dependability were demonstrated by the model's low MAE, MSE, and RMSE values as well as its high Rsquared score, which was attained by utilising LSTM networks' capacity to capture complicated temporal interactions and long-term dependencies. The model's performance was further improved by the thorough preprocessing procedures and the optimised hyperparameters. These results highlight the promise of advanced deep learning methods for financial forecasting, offering insightful information to analysts and investors and opening the door for more studies to improve the predictability and accuracy of financial time series forecasting.

References

[1] Uddin, Rokan, et al. "Multi-Variate Regression Analysis for Stock Market price prediction using Stacked LSTM." 2022 International Conference on Innovations in Science, Engineering and Technology (ICISET). IEEE, 2022.

[2] Waghchaware, Sheetal, and Radhika D. Joshi. "Human activity recognition using 1D-CNN and stacked LSTM." 2022 International Conference on Signal and Information Processing (IConSIP). IEEE, 2022.

[3] Kamepalli, Sujatha, Bandaru Srinivasa Rao, and Kolli Venkata Krishna Kishore. "Multi-class classification and prediction of heart sounds using stacked LSTM to detect heart sound abnormalities." 2022 3rd International Conference for Emerging Technology (INCET). IEEE, 2022.

[4] Prakash, Satya, Anand Singh Jalal, and Pooja Pathak. "Forecasting covid-19 pandemic using prophet, lstm, hybrid gru-lstm, cnn-lstm, bi-lstm and stacked-lstm for india." 2023 6th International Conference on Information Systems and Computer Networks (ISCON). IEEE, 2023.

[5] Abrishami, Soheila, et al. "Enhancing profit by predicting stock prices using deep neural networks." 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI). IEEE, 2019.

[6] Ardeshiri, Reza Rouhi, and Chengbin Ma. "State of charge estimation of lithium-ion battery using deep convolutional stacked bidirectional LSTM." 2021 IEEE 30th International Symposium on Industrial Electronics (ISIE). IEEE, 2021.

[7] Zheng, Haifeng, et al. "A hybrid deep learning model with attention-based conv-LSTM networks for short-term traffic flow prediction." IEEE Transactions on Intelligent Transportation Systems 22.11 (2020)

[8] Gao, Yongqiang, and Zemin Li. "Load balancing aware task offloading in mobile edge computing." 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD). IEEE, 2022.

[9] Xie, Hui, and PengYu Liang. ”Improved TF-LSTM Multi-Step Vehicle Speed Prediction Model Based on LSTM and Attention Mechanism.” 2023 3rd International Conference on Robotics, Automation and Intelligent Control (ICRAIC). IEEE, 2023.